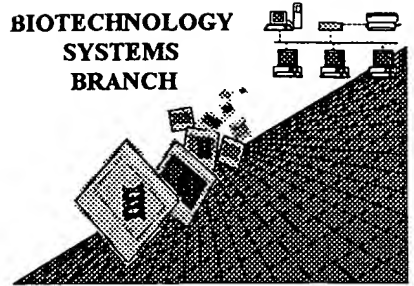


Shukla

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



# **RAW SEQUENCE LISTING** **ERROR REPORT**

**BEST AVAILABLE COPY**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number: 09/357,675  
Art Unit / Team No. : 1632  
Date Processed by STIC: 11/9/99

**THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.**

**PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,**
- 2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY**

**THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.**

**IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:**

**MARK SPENCER 703-308-4212**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/357,675

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 \_\_\_\_\_ Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2 \_\_\_\_\_ Wrapped Aminos      The amino acid number/text at the end of each line "wrapped " down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3 \_\_\_\_\_ Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4 \_\_\_\_\_ Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5 \_\_\_\_\_ Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6 \_\_\_\_\_ Variable Length      Sequence(s) \_\_\_\_\_ contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) feature section that some may be missing.
- 7 \_\_\_\_\_ PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
sequence(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the  
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section  
to the subsequent amino acid sequence.
- 8 \_\_\_\_\_ Skipped Sequences      Sequence(s) \_\_\_\_\_ missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9 \_\_\_\_\_ Skipped Sequences      Sequence(s) \_\_\_\_\_ missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      <210> sequence id number  
<400> sequence id number  
000
- 10 \_\_\_\_\_ Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11 \_\_\_\_\_ Use of <213>Organism      Sequence(s) \_\_\_\_\_ are missing this mandatory field or its response.  
(NEW RULES)
- 12 \_\_\_\_\_ Use of <220>Feature      Sequence(s) \_\_\_\_\_ are missing the <220>Feature and associated headings.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13 \_\_\_\_\_ PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted  
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

*Shukla*

1632

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/357,675

DATE: 11/09/1999  
TIME: 11:27:24

Input Set: I357675.RAW

This Raw Listing contains the General Information  
Section and up to first 5 pages.

Does Not Comply  
Corrected Diskette Needed

*see p. 2, too*

1 <110> APPLICANT: Croce M.D., Carlo M.  
2 <120> TITLE OF INVENTION: Nitrilase Homologs  
3 <130> FILE REFERENCE: CRO01 NP001 Nitrilase Homologs  
4 <140> CURRENT APPLICATION NUMBER: US/09/357,675  
5 <141> CURRENT FILING DATE: 1999-07-20  
6 <150> EARLIER APPLICATION NUMBER: 60/093,350  
7 <151> EARLIER FILING DATE: 1998-07-20  
8 <160> NUMBER OF SEQ ID NOS: 18  
9 <170> SOFTWARE: PatentIn Ver. 2.1  
10 <210> SEQ ID NO 1  
11 <211> LENGTH: 1416  
12 <212> TYPE: DNA  
13 <213> ORGANISM: Homo sapiens  
14 <400> SEQUENCE: 1

*see item 10 on Enr Summary Sheet*

W-->

```
15 gcccactcgc tggggcctnt ctgggtccag accgcctcc ggatcggacc ctgcgaatgg 60
16 ttttggtat atcttcatgt aggacctact cctatcccg tcggccgagg ctgggcttca 120
17 tcaccaggcc tctcacaga ttctgtccc ttctgtgtcc tggactccgg atacctcaac 180
18 tctcagtact ttgtgctcag cccaggccca gagccatggc tatctcctct tctcctgcg 240
19 aactgcccc ggtggctgtg tgccaggtaa catcgacgcc agacaagcaa cagaacttta 300
20 aaacatgtgc tgagctggtt cgagaggctg ccagactggg tgccctgctg gctttcctgc 360
21 ctgaggcatt tgacttcatt gcacgggacc ctgcagagac gctacacctg tctgaaccac 420
22 tgggtgggaa acttttgaa gaatacacc agcttgccag ggaatgtgga ctctggctgt 480
23 ccttgggtgg tttccatgag cgtggccaag actgggagca gactcagaaa atctacaatt 540
24 gtcacgtgct gctgaacagc aaaggggcag tagtggccac ttacaggaag acacatctgt 600
25 gtgacgtaga gattccaggg caggggcta tgtgtgaaag caactctacc atgctgggc 660
26 ccagtcttga gtcacctgtc agcacaccag caggcaagat tgggtctagt gtctgctatg 720
27 acatgcggtt cctgaactc tctctggcat tggctcaagc tggagcagag atacttacct 780
28 atccttcagc ttttgatcc attacaggcc cagccactg ggaggtgttg ctgcggggccc 840
29 gtgctatcga aaccagtgct tatgtagtgg cagcagcaca gtgtggacgc caccatgaga 900
30 agagagcaag ttatggccac agcatggtgg tagaccctg gggaacagtg gtggcccgc 960
31 gctctgaggg gccaggctc tgccctgccc gaatagacct caactatctg cgacagttgc 1020
32 gccgacacct gctgtgttc cagcaccgca ggccctgacct ctatggcaat ctgggtcacc 1080
33 cactgtctta agacttgact tctgtgagtt tagacctgcc cctcccaccc ccacctgcc 1140
34 actatgagct agtgtcatg tgacttgagg gcaggatcca ggcacagctc ccctcacttg 1200
35 gagaaccttg actctcttga tggaaacacag atgggctgct tgggaaagaa actttcacct 1260
36 gagcttcacc tgaggtcaga ctgcagtttc agaaaggtgg aattttatat agtcattgtt 1320
37 tatttcatgg aaactgaagt tctgctgagg gctgagcagc actggcattg aaaaatataa 1380
38 taatcataaa gtcaaaaaaa aaaaaaaaaa aaaaaa 1416
```

39 <210> SEQ ID NO 2  
40 <211> LENGTH: 23  
41 <212> TYPE: DNA  
42 <213> ORGANISM: Homo sapiens  
43 <400> SEQUENCE: 2  
44 tctgaaactg cagtctgacc tca

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/357,675

DATE: 11/09/1999  
TIME: 11:27:24

Input Set: I357675.RAW

45 <210> SEQ ID NO 3  
46 <211> LENGTH: 21  
47 <212> TYPE: DNA  
48 <213> ORGANISM: Homo sapiens  
49 <400> SEQUENCE: 3  
50 caggcacagc tcccctcact t 21  
51 <210> SEQ ID NO 4  
52 <211> LENGTH: 20  
53 <212> TYPE: DNA  
54 <213> ORGANISM: Homo sapiens  
55 <400> SEQUENCE: 4  
W--> 56 ~~gtgtgttcggt gncaygttgt~~ *idem 10* 20  
57 <210> SEQ ID NO 5  
58 <211> LENGTH: 26  
59 <212> TYPE: DNA  
60 <213> ORGANISM: Homo sapiens  
61 <400> SEQUENCE: 5  
W--> 62 ~~acrtgtacrt gyttnacngt ytgngc~~ *idem 10* 26  
63 <210> SEQ ID NO 6  
64 <211> LENGTH: 21  
65 <212> TYPE: DNA  
66 <213> ORGANISM: Drosophila melanogaster  
67 <400> SEQUENCE: 6  
68 gcgcctttgt ggcctcgact g 21  
69 <210> SEQ ID NO 7  
70 <211> LENGTH: 21  
71 <212> TYPE: DNA  
72 <213> ORGANISM: Drosophila melanogaster  
73 <400> SEQUENCE: 7  
74 cgggtggcgga agttgtctgg t 21  
75 <210> SEQ ID NO 8  
76 <211> LENGTH: 20  
77 <212> TYPE: DNA  
78 <213> ORGANISM: Caenorhabditis elegans  
79 <400> SEQUENCE: 8  
80 gtggcggctg ctcaaactgg 20  
81 <210> SEQ ID NO 9  
82 <211> LENGTH: 21  
83 <212> TYPE: DNA  
84 <213> ORGANISM: Caenorhabditis elegans  
85 <400> SEQUENCE: 9  
86 tcgcgacgat gaacaagtgc g 21  
87 <210> SEQ ID NO 10  
88 <211> LENGTH: 19  
89 <212> TYPE: DNA  
90 <213> ORGANISM: Homo sapiens  
91 <400> SEQUENCE: 10  
92 gccctccgga tcggaccct 19  
93 <210> SEQ ID NO 11  
94 <211> LENGTH: 20

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/357,675DATE: 11/09/1999  
TIME: 11:27:24

Input Set: I357675.RAW

95	<212> TYPE: DNA	
96	<213> ORGANISM: Homo sapiens	
97	<400> SEQUENCE: 11	
98	gacctactcc ctatcccgtc	20
99	<210> SEQ ID NO 12	
100	<211> LENGTH: 21	
101	<212> TYPE: DNA	
102	<213> ORGANISM: Homo sapiens	
103	<400> SEQUENCE: 12	
104	gctgcgaagt gcacagctaa g	21
105	<210> SEQ ID NO 13	
106	<211> LENGTH: 24	
107	<212> TYPE: DNA	
108	<213> ORGANISM: Homo sapiens	
109	<400> SEQUENCE: 13	
110	aaactgaagc ctctttcctc tgac	24
111	<210> SEQ ID NO 14	
112	<211> LENGTH: 20	
113	<212> TYPE: DNA	
114	<213> ORGANISM: Homo sapiens	
115	<400> SEQUENCE: 14	
116	tgggcttcat caccaggcct	20
117	<210> SEQ ID NO 15	
118	<211> LENGTH: 22	
119	<212> TYPE: DNA	
120	<213> ORGANISM: Homo sapiens	
121	<400> SEQUENCE: 15	
122	ctgggctgag cacaaagtac tg	22
123	<210> SEQ ID NO 16	
124	<211> LENGTH: 21	
125	<212> TYPE: DNA	
126	<213> ORGANISM: Homo sapiens	
127	<400> SEQUENCE: 16	
128	gcttgtctgg cgtcgatgtt a	21
129	<210> SEQ ID NO 17	
130	<211> LENGTH: 36	
131	<212> TYPE: DNA	
132	<213> ORGANISM: Homo sapiens	
133	<400> SEQUENCE: 17	
134	tgacgtcgac atatgtcaac tctagttaat accacg	36
135	<210> SEQ ID NO 18	
136	<211> LENGTH: 25	
137	<212> TYPE: DNA	
138	<213> ORGANISM: Homo sapiens	
139	<400> SEQUENCE: 18	
140	tgggtacctc gactagctta tgtcc	25

VERIFICATION SUMMARY  
PATENT APPLICATION US/09/357,675DATE: 11/09/1999  
TIME: 11:27:24

Input Set: I357675.RAW

Line	Error/Warning	Original Text
15	W "N" or "Xaa" used: Feature required	gcccactcgc tgcggcctnt ctggctccag accgcct
56	W "N" or "Xaa" used: Feature required	gtngtnccng gncaygtngt
62	W "N" or "Xaa" used: Feature required	acrtgnacrt gyttnacngt ytgngc